KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and Applied Economics

The small firm anomaly: US and international evidence

Lieven De Moor and Piet Sercu

DEPARTMENT OF ACCOUNTANCY, FINANCE AND INSURANCE (AFI)

AFI 0609

The Small Firm Anomaly: US and International Evidence*

Lieven De Moor†
European University College Brussels
Katholieke Universiteit Leuven

Piet Sercu‡
Katholieke Universiteit Leuven

First version: 29/03/2004
Current version: 17/10/2006

Abstract The results of a CAPM test are sensitive to aspects related to the weight one gives to small, low-visibility stocks when constructing the portfolios whose returns serve as left- and right-hand-side variables. It turns out to be the result of a marketwide factor rather than a stock characteristic. To fit the observed returns it suffices to redesign the size and book-to-market factor portfolios into two factor portfolios each, one for the smallest or highest book-to-market stocks relative to other stocks, and one for moderately small or high book-to-market stocks versus larger or growth companies. This alternative 6-factor model does a better job in pricing stocks, both in the US and internationally, than Carhart's 4-factor CAPM with factor portfolios designed following Fama and French (1992, 1993, 1995, 1996a, 1996b, 1998, 2000), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999). The fact that we can resolve mispricing by adding factors rather than characteristics, rules out data

---

* An early precursor paper by the same authors, entitled *CAPM Tests and Alternative Factor Portfolio Composition: Getting the Alphas Right*, appeared as DTEW Research Report 0521, KULeuven, 60p. and has been presented at AFFI 2004 Paris and EFA 2005 Moscow

† Corresponding author: Lieven De Moor, Stormstraat 2, B-1000 Brussels, Belgium. Tel.: +32 498 61 64 63. Fax.: +32 2 210 12 22. E-mail: lieven.demoor@ehsal.be

‡ Piet Sercu, Naamsestraat 69, B-3000 Leuven, Belgium. Tel.: +32 16 32 67 56. Fax.: +32 16 32 67 32. E-mail: piet.sercu@econ.kuleuven.be

The authors thank Marno Verbeek, Geert Dhaene, Constant Beckers, Stefan Duchateau, Cynthia Van Hulle, Marcello Pericoli, Mark Van Achter, Adam Szyszka, Kevin Aretz, Pim van Vliet and seminar participants at the Katholieke Universiteit Leuven, the French Finance Association (Paris, 2004) and the European Finance Association (Moscow, 2005) for helpful discussions and valuable comments on earlier drafts of this paper. Lieven De Moor gratefully acknowledges financial support from the Fund for Scientific Research - Flanders (Belgium).

problems as an explanation, and information asymmetries. Thin trading bias in the beta is also rejected as the source of extra returns. Liquidity remains a serious possible candidate, as is the hypothesis of extra downside risk for small firms.

JEL codes: G12, G15.
Keywords: Small Firms, CAPM, SMB, HML, WML.

1. Introduction

The empirical anomalies that emerged from CAPM tests, such as the size, book-to-market and momentum effects (Banz, 1981; Stattman, 1980 and Rosenberg, Reid and Lanstein, 1985; Jegadeesh and Titman, 1993), have quickly been incorporated into generalized asset pricing models. Empirical work by e.g. Fama and French (1992, 1993, 1995, 1996a, 1996b, 1998, 2000), Carhart (1997) or Rouwenhorst (1999) reveals that these additional factor portfolios significantly improve the model's ability to capture the cross-sectional variation of stock returns, both within the US and internationally. The purpose of this paper is to extend these tests to a dataset that has a wide coverage both across the size spectrum and across countries. We find that both in the US and internationally the ten percent smallest stocks do not fit the standard models.

The question naturally arises where these unusually large returns come from. One possible explanation is information asymmetries. These are more of a problem the smaller the stocks: analyst following and press coverage are positively related to size. A second possible explanation is liquidity. Amihud and Mendelson (1986), for instance, document that liquidity, as measured by the bid-ask spread, subsumes the size effect in returns from equities. Third, the small companies have more downside risk because they are more exposed to large negative shocks (Chan and Chen, 1988, 1991; Chen and Zhang. 1998). Fourth, despite our best efforts we might have missed some data errors in Datastream; shifts in the decimal sign for a stock price, for instance, bias the average returns upwards as they generate spurious +900%, -90% or -90%,+900% episodes.

We can get a first insight into the likely cause(s) of the high returns by using the distinction between a characteristic and a factor (Daniel *et al.*, 1997). In the case of a characteristic, returns are explained by, for example, size or leverage or dividend yield. Such an attribute of the company is very different from a factor: a factor is always time-varying, it is active for many or all stocks, and it is the company-specific sensitivity to the factor—a covariance—that explains expected returns.

Data errors cannot be picked up by a factor since, being random, they do not co-vary with a market-wide variable. (One exception must be made for data errors caused by stale prices (thin trading): we control for this possible explanation separately.) Liquidity, when instrumented by the bid-ask spread as in Amihud and Mendelson (1986), is more like a characteristic. But liquidity could be modeled as a factor too if it is directly linked to turnover instead of spread. Variations over time in a company's turnover are clearly correlated across stocks and low liquidity when all stocks suffer from a thin market is much worse than idiosyncratic episodes of low depth. In this sense, liquidity may be a factor. Information asymmetry is less likely to be a factor: there is little obvious variation over time, and even less variation that goes together with market-wide information problems.[1] Sensitivity to negative shocks in the business cycle, lastly, corresponds a factor; in Chen *et al.* (1985), the factor is, for instance, correlated to the net business formation.

Directly testing some of the candidate factors, like net business formation or the market-averaged bid-ask spread, is difficult since most countries do not have easily available data on them. To get at least part of the insights we resort to a shortcut. If the explanation is a priced factor, it can be picked up by a portfolio of assets, provided that the return-generating process is sufficiently close to linearity and the residual returns nearly independent. Thus, if we can identify portfolios that resolve the mispricing via their covariances, we narrow down the list of explanations to factors. This, then, would already rule out data errors and information asymmetries. Our evidence indeed is that two extra portfolios suffice to resolve most of the pricing problems in the US. Better, these extra portfolios also work internationally, that is, outside the market that suggested the specifications of the extra factor portfolios. It turns out that the required factor portfolios are one extra return spread for size and one for book-to-market. The resulting generalized model provides a risk-return relation that outperforms national and global one- or Carhart's 4-factor CAPMs (with market, size, book-to-market, and momentum portfolios), and the nested version of an international CAPM and Carhart's 4-factor model.

The structure of this paper is as follows. Section 2 describes the dataset. Section 3 focuses on the US market, the subject of most of the extant research. Our starting point is a replication of the Fama and French (1993, 1996) tests. When following the Fama and French (1993) procedure as closely as possible *re* data coverage we do find similar results as the original study, despite the different period (1980-1993 rather than 1963-1993). We then gradually modify the procedure, and notably increase the data coverage and the room given to small stocks. The result is large positive alphas for the lowest size decile and U and J shapes across the board. Thus, we may need to add not just momentum but also an extra

[1] True, asymmetries are often measured by bid-ask spreads, and these do co-vary across stocks. But spreads are also driven by liquidity, where there is a very clear market-wide factor; so co-variation in spreads is more likely to reflect a liquidity factor than an information factor.

small-firm and book-to-market factor to the original Fama and French (1993) trio (market, SMB and HML). Our tests of this candidate factor specification in the US market, in Section 4, reveal that the alternative 6-factor model does explain various style portfolio returns, whether stratified across one or two styles and whether separated from the factor portfolio data or not. In Section 5, then, we venture beyond the US borders and successfully test our alternative asset pricing model against competing models, using various style portfolios (size, book-to-market, and momentum). Section 6 concludes.

2. The Dataset

Our aim was to create an international complete and clean equity list, offering maximal coverage within and across countries, minimal data errors and minimal duplications. The Datastream Research lists are quite international and claim to contain all quotes on (all) the exchange(s) of the specified country. This means that a large number of small firms are included. Unfortunately these lists also contain small, illiquid and penny stocks, as well as secondary or tertiary listings; in addition, it suffers from survivorship bias. The Datastream Dead lists are the "dead-version" of the Datastream Research lists, containing all delisted stocks on the exchanges of the specified country. We merged both lists, cleaned the merged list for unwanted assets and cleaned the time series for bad data (see Section 2.1). As our aim is to compose an international database, we chose countries on the basis of data availability taking into account coverage within and across regions: North America (Canada, United States), Latin America (Argentina, Brazil, Chile, Colombia, Mexico, Peru), Japan, Asia-ex-Japan (China, Hong Kong, India, Indonesia, Malaysia, Philippines, Singapore, South Korea, Taiwan, Thailand), Euro-in countries (Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxemburg, Netherlands, Portugal, Spain), Euro-out countries (Denmark, Greece, Norway, Sweden, UK), Switzerland, Australasia (Australia, New Zealand) and South-Africa.

2.1 Cleaning the Data List:

We deleted:

- Dual listings within and across Exchanges (e.g. ADR's, GDR's, identical shares), preferred shares, warrants, certificates, shares from the same company but with different voting rights.
- Error shares (e.g. shares with no name, one-day shares).

- Special sectors (e.g. funds, trusts, investment companies, financial holding companies) i.e. shares that duplicate information on individual companies.

2.2 Cleaning the Time Series:

The resulting equity list contains 44318 unbalanced time series of dollar returns, not prices or local-currency returns. We applied a filter that eliminates small, illiquid and penny stocks. Penny stocks have a larger probability to contain errors. They are often fallen stocks which are highly speculative and illiquid. Small companies also have limited liquidity, can be subject to high price pressure or price manipulation, and often represent too little value to warrant attention. In practice this means that an end-of-month price formation of a stock with a market capitalization smaller than $10,000,000 or a monthly trading volume smaller than $100,000 or a price smaller than $1, are eliminated. If trading volume information is not available, we considered an unchanged monthly price (in local currency) as a sign of low trading volume and unreliable price formation for that month and hence both returns based on this price are eliminated. Lastly, we eliminated all stock quotes corresponding to a negative book-to-market value.

After applying this filter we still encountered a few high-return errors. Apparently Datastream contains some returns that are simply too good to be true and can be very influential for regression results. The few high-return errors we encountered were: (1) decimal-sign shifting; (2) anomalously low first price of a series (probably theoretical or illiquid); (3) high reported return not corresponding to a similar change in the market capitalization, price or to a huge dividend payout; (4) data reported before actual introduction date or after the actual delisting date; (5) obvious typos; (6) wrongly handled equity offerings. All these were treated as missing observations.

2.3 Descriptive Statistics

36% of the stocks do not have any book values in the database. The median market value of those stocks is $60,930,000 whereas the median market value of the other 64% is $135,060,000. So, stocks without book values are primarily smaller stocks and take a significant part of the database. The stock list contains 44318 unbalanced time series with the geographical and sectoral distribution shown in Figures 1 to 5.

3. The Fama and French (1993) Model and the Small-Firm Anomaly

Our first test design tries to be as close as possible to Fama and French (1993). We then test the robustness to modified designs. Some modifications are inspired by data availability outside the US, but the main change is the increased room for smaller stocks. We first review the original Fama and French (1993) procedure.

In Fama and French (1993) monthly dollar returns on 25 portfolios of size-and-book-to-market-sorted stocks are regressed on three factor portfolios: the market portfolio, the size factor and the book-to-market factor, i.e.

$$R_i - r_f = \alpha_i + \beta_i (R_m - r_f) + \gamma_i SMB + \delta_i HML + \varepsilon_i \,. \qquad (1)$$

At the end of June of each year, stocks are allocated to either of two groups (small or big, denoted S or B) depending on whether their early-June market cap is below or above the median market equity for NYSE stocks.[2] All stocks are also allocated, via an independent second sort, into one of three book-to-market (B/M) equity groups (low, medium, or high, denoted L, M, or H); the watershed values are the 30th and 70th percentile values of B/M-ranked NYSE stocks. Note that Fama and French (1993) only use stocks for which both market values and book values are available, which may not be an innocent restriction on the sample.

They then proceed as follows. For the purpose of constructing the factors, six size-B/M portfolios are then defined as the six intersections of the two size groups and the three B/M groups. These six intersections are labeled S/L, S/M, S/H, B/L, B/M, and B/H. Value-weighted monthly returns on the six portfolios are calculated from July till June next year. For each month, the size factor SMB is computed as the difference between the returns on small stocks (the average of the returns on the three small-stock portfolios, S/L, S/M and S/H) and big stocks (the average returns on the three big-stock portfolios, B/L, B/M and B/H). The book-to-market factor HML is the difference between, on the one hand, the average of the returns on the two high B/M portfolios (S/H and B/H) and, on the other, the average of the returns on the two low B/M portfolios (S/L and B/L). Note that the returns are value-weighted within each of the six size-B/M portfolios, while for the calculation of SMB and HML, equally weighted averages are taken across the three S/. or B/. portfolios.

[2] Fama and French (1993) do use only the NYSE stocks to set allocation breakpoints for both size and book-to-market, not the median of all stocks (including Amex and NASDAQ). The reason for this is not stated explicitly.

For the purpose of generating test portfolios (that is, portfolios whose returns need to be explained by the factors), Fama and French (1993) form 25 size-B/M portfolios following the same procedure as for the six size-B/M portfolios underlying SMB and HML, except that quintile breakpoints for size and B/M for NYSE stocks are used to allocate all stocks to the portfolios rather than the median or the 30th and 70th percentile values. Fama and French (1993) discard negative-B/M firms when calculating the breakpoints or forming size-B/M test portfolios. For the reader's convenience, Table 1 reproduces the alphas obtained in the original Fama and French (1996a) study.

To set the stage, we replicate the above Fama and French (1993) test on our database (Table 2). The overpricing (or return shortfall) that, in Fama and French (1996a), occurred for the small, growth stocks seems to have shifted up one class, into the second size quintile, possibly because our database contains more smaller stocks than the Fama and French (1996a) database. In addition, the return anomalies for the book-to-market stocks (the rightmost column) have become more pronounced, both algebraically and statistically. There may also be evidence of what looks like interactions: the extreme size/book-to-market combinations show most mispricing, with the corner cases on the main diagonal being overpriced and those on the secondary diagonal underpriced. Still, the differences are not massive.[3]

Table 3 is the result of a test of the Fama and French (1993) model that uses a different design. Table 4 is the roadmap that lists the differences in design and traces changes between Table 2 (our replication of the Fama and French (1993) procedures on our data) and Table 3. The choices *re* the research period, the risk-free rate and the market return have only a minor impact on the number of significant alphas. The first major source of differences is monthly rebalancing. Switching from yearly updated test- and/or factor portfolios to their monthly updated variants boosts the significant alphas both in numbers and values. As there is no

[3] As expected, our portfolio returns do not exhibit significant autocorrelation. But the returns do exhibit conditional heteroskedasticity over time. Following Fama and French (1993) we initially ignore this, but towards the end we do shift to a heteroskedasticity-consistent covariance matrix of OLS/SUR. It appears that the heteroskedasticity-consistent covariance matrix leads to fewer significant alphas than the plain OLS ones, but the difference is never very pronounced. With identical regressors across equations and no cross-equation restrictions, Seemingly Unrelated Regression (SUR) provides the same estimates and standard errors as OLS. But the weighting matrix we use is White's heteroskedasticity-consistent covariance matrix, so that the significance statements are robust to both heteroskedasticity (over time or across stocks) and contemporaneous correlation of unknown form.

obvious explanation why this should be so, except that frequent rebalancing may increase power, we keep on using monthly updated portfolios in the tests below. We treat it as an anomaly or at least an issue of robustness that should be resolved.

The most important source of differences between Table 2 and Table 3, however, is the weight given to small stocks. The Fama and French (1993) procedure gives relatively little room to small firms in three respects:

- It discards stocks for which either book values or market values are missing, a restriction that tends to eliminate mostly small companies. Thus, the standard SMB and HML may overlook part of a small-firm effect. Simultaneously, any such deficiency in the factors may never show up because the companies most affected by the potentially missed factor are absent on the left-hand side too.
- The assignment of stocks to factor portfolios or test portfolios is based on NYSE percentile values even though the database also includes Amex and NASDAQ stocks. This results in size groups with more firms in the smaller categories and, likewise, book-to-market groups with more stocks in the growth or low book-to-market category.
- Value-weighting: while the portfolio-theory logic underlying the CAPM dictates value weights as far as the market portfolio is concerned, there is no such theoretical basis for the size and book-to-market factors. One effect of value weighting is that the Fama and French (1993) S(mall) portfolio, even though it contains all below-median stocks, is dominated by the comparatively larger ones, those close to the median size.[4] Since, in addition, the median is the NYSE one, the value-weighted S(mall) portfolio may be more of a mid-cap portfolio than the small-cap one like its name would suggest.

While there is, of course, nothing a priori wrong with all this, a robustness check seems useful. For these reasons we use (a) "broad-based" portfolios and breakpoints (based on all stocks) instead of "narrow-based" portfolios and breakpoints from stocks that have both book values and market values; (b) breakpoints based on all stocks instead of NYSE stocks only and; (c) equally weighted portfolios for factors other than the market. We do this broadening on left-hand side ("test portfolios") and on the right-hand side ("factor portfolios"). We find, in Table 4, that the broadened factor portfolios (and breakpoints) are more capable of pricing

[4] The fact that S is actually computed from the (value-weighted) returns of three size/book-to-market intersections, only partly mitigates this effect, because the relation between size and book-to-market is far from perfect.

unmanaged size-book-to-market portfolios: the number of rejections drops, from 16 to 7. This strongly suggests that the new factor specification is a step in the right direction: the Fama and French (1993) factors, by restricting the coverage to stocks with both a known market value and a known book-to-market value, miss too many of the smaller firms. However, the use of broadened *test* portfolios worsens the fit again as the number of significant alphas rises from 7 to 14. This strongly suggests that broadened test portfolios are more powerful in the sense of providing more rejections. For these reasons we continue to apply broad-based factor and test portfolios in the following tests.

The large number of significant abnormal returns is not the only anomalous result in Table 3. In addition, the typical rejected alpha is about 0.5% per month or more, which is worse than the kind of numbers Fama and French (1996a) obtain. Lastly, there are manifest patterns in the alphas. First, within each and every row there is a U shape in the alphas. Second, within each column there is a J shape, with the small-firm quintile always providing a strongly positive abnormal returns, the second quintile a strongly negative one, followed by gradually improving returns for higher-size quintiles. It is, we think, fair to say that the three-factor model does not span our returns very well and that size seems to be part of the problem.

4. Identifying the Missing Factors: US Data

In view of the momentum-related anomalies that came to light after the publication of Fama and French (1993), a momentum factor is added to the model, following the Rouwenhorst (1999) version of Jegadeesh and Titman (1993). This does not eliminate the small-stock problem, though. In this section we apply a generalized Fama and French (1993) model that takes care of most of the anomalies we just noted, without having to bring in new non-return data like bid-ask spreads.

This section is organized as follows. We start, in Section 4.1, with a look at mean returns on decile portfolios, one set per risk dimension. Even though the sorting is one-dimensional and the returns are not risk-adjusted, we find back the U's and J's we observed in the alphas of the previous section. A closer scrutiny of one-dimensional decile portfolios provides ideas on how to define the additional factors (Section 4.2). The resulting alternative 6-factor model in its full version is then successfully tested against Carhart's 4-factor model (Section 4.3). The obvious risk, in this approach, is that we might be over-fitting a specific dataset;

however, bear in mind that the resulting model is tested also on international data (Section 5), where it appears to hold well, too.

4.1 One-dimensionally sorted portfolios: the role of size revisited

In this section we look at returns from decile portfolios of stocks sorted along one dimension at the time.

Figures 6 to 8 show average returns for ten decile portfolios sorted by size, book-to-market (B/M) and momentum, respectively. For further reference we make a comment for each type of stratification.

First, the main size effect is found in the first and to some extent also in the second decile, which provides higher average returns. In deciles 3-10, in contrast, there is a weak premium for larger sizes.

Second, we note a book-to-market effect is more monotone positive, but S-shaped rather than linear. There is a mild return-shortfall effect in deciles 1 and 2 (growth firms earning moderately lower returns), which then flattens out; and as of decile 7, "value" firms earn increasingly higher premia.

Third, an S-pattern is also present in the momentum factor, with strong losers going on earning clearly lower returns, strong winners continuing their upward trend, and flat returns for a wide midrange (deciles 4-8).

Common to the three schedules is the nonlinearity. These patterns raise the possibility that the tradition of capturing the size factor (or book-to-market or momentum factor) by just one number, the difference between a "hi" and a "lo" portfolio return, may be too simplistic. True, this inference is indicative only. For one thing, in theory the stocks' sensitivities to the factors could be sufficiently nonlinear in the quantile's order i to pick up the apparent nonlinearity. Second, the sort is one-dimensional; in theory the omitted other risk factors could still be responsible for what here seems to be a nonlinearity. Still, recall that we obtained very similar conclusions from the alphas in the previous section (Table 3), where exposures to factors were used rather than quantile membership and where two dimensions of non-market risk were considered simultaneously. In the next subsection we identify factor portfolios that get the alphas of unmanaged funds as close to zero as possible.

4.2 In search of optimal factor portfolios

The conjecture behind the rest of this paper is that the apparent mispricing in Section 3 may be resolved by using, in every risk dimension, two return differentials rather than one. We compare the performance of this alternative model, with its six factors, to competing models. In Section 5 we then test the approach largely out-of-sample, namely on international data.

4.2.1 A second size factor

The average monthly dollar returns of the size deciles for the period 1980-1999 were already shown in Figure 6, which revealed a large average return for the first-decile (smallest) stocks and a slightly higher average return for the largest stocks compared to the middle deciles. This last observation is in line with Fama and French (1992), who find evidence that the size premium in the US has become weaker in recent years. In fact, for 1980-1990 they document a negative size premium. Eun, Huang and Lai (2003) likewise find that, recently, the mean return is somewhat higher for large-cap funds than for small-cap funds in the US market. But from Figure 6 and the existing literature, it seems that there still exists a strong small-firm effect for the first-decile stocks in the US that is missed by databases that are too selective: only when we go beyond our first- and second-decile stocks we see an inverted small-firm effect in the US. All this was about raw returns linked to decile membership, not returns risk-corrected via regression exposure coefficients.

In search of a portfolio proxy for the unidentified size-related factor(s), we hypothesize two kinds of size risk: (*a*) the regular size factor like in Fama and French (1993) which holds for all stocks but the smallest; and (*b*) the risk inherent to the smallest stocks that cannot be accounted for by neither beta risk nor the regular Fama and French (1993) size risk. Since Fama and French (1993) already coined the label *small* for their not-so-small stock portfolio, we chose the label *micro stocks* for our new factor, even though by many countries' standards these micro stocks are still quite sizable. We get the best results for purely size-sorted deciles with a micro-stock risk factor (mSMB) defined as the return on a zero-investment portfolio that is long the first decile and short deciles 2 to 10, and a regular size risk factor, defined as a zero-investment portfolio that is long in stocks from deciles 2 and 3 and short stocks from deciles 6 to 9.

4.2.2 A second book-to-market factor

The average monthly dollar returns for the book-to-market-decile portfolios for the period 1980-1999 were already shown in Figure 7. Recall that we saw a monotone positive but S-shaped schedule where the highest-book-to-market decile really sticks out, which might indicate an extra book-to-market risk.

In search of a portfolio proxy for the unidentified book-to-market-related factor(s), we hypothesize two kinds of book-to-market risk: (*i*) extreme book-to-market risk, i.e. the risk inherent to the highest B/M stocks that cannot be accounted for by beta risk nor normal book-to-market risk; (*ii*) normal book-to-market risk in the spirit of Fama and French (1993) but redefined to reduce overlap with extreme book-to-market risk. We get the best results for purely book-to-market-sorted deciles with an *eHML* factor reflecting extreme risk defined as the return on a zero-investment portfolio that is long the highest B/M-decile stocks and short all other B/M deciles; and a *rHML* risk factor reflecting the regular book-to-market risk, measured as the return on a zero-investment portfolio that is long the value stocks in B/M deciles 8 to 10 and short the growth stocks B/M deciles 1 and 2.

4.2.3 A modified momentum factor

The average monthly dollar returns of the momentum deciles for the period 1980-1999 were already shown in Figure 8. We described the plot as a mildly S-shaped rise. The apparent nonlinearity may still be picked up by the difference between decile membership and exposure, or by beta or other risks.

When we combine the standard momentum portfolio with the standard size and book-to-market portfolio, i.e. Carhart's 4-factor model, two portfolios are still mispriced (Table 5, to be discussed below). It turns out that no second momentum portfolio is needed to mend this. We get the best results by redefine *WML* as the difference between returns from the 10% winners and the 20% losers.

In the next section we combine these alternative risk factors to one multi-factor model and test it formally on two-dimensionally sorted portfolios.

4.3 Tests of the modified factor specification

In this section we combine the alternative size-, book-to-market- and momentum risk factors into one alternative 6-factor model. We show that, in pricing different kinds of unmanaged portfolios, this model does a better job than Carhart's 4-factor model with factor portfolios like in Fama and French (1993, 1995, 1996a, 1996b), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999). Carhart's 4-factor model is:

$$R_i - r_f = \alpha_i + \beta_i (R_m - r_f) + \gamma_i SMB + \delta_i HML + \phi_i WML + \varepsilon_i \, , \tag{2}$$

where *SMB* (small minus big) is the size factor portfolio, *viz.* a zero-investment portfolio that is long the 50% smallest stocks and short the 50% largest stocks; *HML* (high minus low) is the book-to-market factor portfolio, a zero-investment portfolio that is long the 30% highest B/M stocks and short the 30% lowest B/M stocks; and *WML* (winner minus loser) is the momentum factor portfolio, a zero-investment portfolio that is long the 30% top past performers (winners) and short the 30% lowest past performers (losers). All portfolios are equally weighted and updated monthly.

Our alternative 6-factor model uses the same factor portfolios (size, book-to-market and momentum) albeit with a slightly modified definition, plus the two extra factors for extreme size and book-to-market risks. Combining the alternative factor portfolios from the preceding sections into one model gives the following alternative 6-factor model:

$$\begin{aligned} R_i - r_f = {} & \alpha_i + \beta_i (R_m - r_f) + \gamma_i mSMB + \delta_i rSMB + \phi_i eHML \\ & + \varphi_i rHML + \theta_i WML + \varepsilon_i \, , \end{aligned} \tag{3}$$

with the factors as defined in Section 4.2.

4.3.1 The pricing of one- and two-dimensional test portfolios

We demonstrate that the alternative 6-factor model is a better model to price unmanaged one-dimensional test portfolios—that is, stocks sorted on either size, book-to-market or momentum—than Carhart's 4-factor model. To give Carhart's model every chance it deserves, we introduce it in four different implementations. The first is the one we described below Equation 2. The variants are the following:

- We showed in Section 3 that switching from yearly updated factor portfolios to monthly updated factor portfolio (SMB and HML in Table 4) generated more significant alphas (13 against 5). For this reason we also use yearly updated—as an alternative to of monthly updated—factor portfolios (SMB and HML).
- It is also possible that the beta coefficient of small stocks is underestimated because of thin trading associated with small stocks. Thin trading introduces a bias towards zero in

the contemporal covariance of small stocks with the market return. We can substantially mitigate this problem by including the two leads and lags of the market return as additional factors. This way we enlarge the market return window with which small stocks can covary as in Dimson (1979).[5]

- Last, we work with Carhart's model that uses the original risk factors (SMB and HML) from Fama and French (1993), updated to December 1999.[6]

In Table 5 the estimated alphas of the alternative 6-factor model are always insignificantly different from zero and Wald's test cannot reject the null hypothesis of zero alphas. Under Carhart's 4-factor model with monthly or yearly updated factor portfolios (SMB and HML), in contract, there are six or four alphas significantly different from zero and Wald's tests always reject the null hypothesis of zero alphas. The significant alphas are clearly not eliminated by adding two leads and lags of the market risk factor or by using the original Fama and French (1993) risk factors. We therefore conclude that neither the underestimated-beta argument nor the original Fama and French (1993) risk factors are able to make Carhart's 4-factor model fit the broader sample. We conclude that the alternative 6-factor model is a better model in pricing unmanaged size-, book-to-market- or momentum sorted portfolios.

Table 6 shows that for each of the three two-dimensional (5x5) test portfolios the alternative 6-factor model generates far fewer significant alphas, and lower χ^2 statistics, than Carhart's 4-factor model. For the size-book-to-market-sorted test portfolios that fared so badly in the Fama and French (1993) tests of Section 3, the number of rejections falls from 7 to just 1. For the size-momentum and momentum-book-to-market test portfolios we go from 4 to 1 or zero. We conclude that the alternative 6-factor model does a better job in pricing unmanaged size-book-to-market, size-momentum and momentum-book-to-market portfolios than Carhart's model.

---

[5] Dimson (1979) estimates market sensitivities (betas) in the presence of thin trading via a multiple regression that includes leads and lags of the market return–leads because part of today's true market return will show up tomorrow only because some stocks do not trade today, and lags because for a stock that did not trade yesterday, today's reported return is partly explained by yesterday's true market return.

[6] Downloaded from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

4.3.2 Out-of-sample tests

Fama and French (1995) point out that spurious common variation might be induced when the regressor portfolios *SMB* and *HML* are constructed from the same stocks as the regressand test portfolios.[7] To avoid this, they provide a test where the stocks in the left-hand-side portfolios are different from those on the right-hand side. Specifically, they split the data into two equal groups. One group provides the dependent value-weighted size-B/M test portfolios for the time-series regressions. The other is used to form explanatory factor portfolio returns.

We proceed similarly. Stocks are ranked alphabetically; the odd-numbered are used to calculate the left-side test portfolios, from the even-numbered stocks we calculate the right-side risk factors.

This procedure produces reassuringly similar alpha estimates and *t*-statistics. Thus, spurious correlation does not seem to have been behind our earlier good results. This is in line with Fama and French (1995). Table 6 also shows the importance of sample size: in Carhart's model, the rejections are down from seven to five; still, the alternative 6-factor model comes out with just one.

4.4 Conclusion

The alternative 6-factor model provides a significantly improved version relative to Fama and French's 3- or Carhart's 4-factor CAPM with factor portfolios of Fama and French (1993, 1995, 1996a, 1996b), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999). Our model produces estimated alphas closer to zero for one-dimensional size, book-to-market and momentum portfolios and two-dimensional size-book-to-market, size-momentum and momentum-book-to-market portfolios. However, the evidence so far bears on the US market only, and the factor portfolios were hand-picked to fit this dataset. In the next section, we accordingly test whether the alternative factor portfolios keep on producing estimated alphas close to zero in an international setting.

[7] Fama and French (1995) investigate size and book-to-market factors, a momentum factor is not included.

5. International Validation

Recall that Fama and French (1993) calculated the size- and book-to-market-decile breakpoints on the NYSE stocks only, and used these to catalogue all US stocks, including Amex and NASDAQ stocks. One possible issue is that this procedure is difficult to implement in an international setting. When Fama and French (1998) investigate value versus growth effects in an international setting, they abandon this procedure and calculate the decile breakpoints from all stocks.

They proceed by calculating their size and book-to-market factor portfolios (*SMB* and *HML*) for each country separately. The global factor portfolios are then constructed as averages of these national factor portfolios, weighted according to the MSCI country weights. This surely avoids the risk that, say, the Big portfolio becomes very much a US affair. However, in some countries the range of corporate size or B/M is quite narrow: many small Western countries have no really big firms, and some emerging markets specialize in one sector, thus reflecting the rather similar size or book-to-market figures that are typical for that sector. In short, one issue is whether classification into, say, the Big or Small buckets should be done country by country or via one global list.

Another issue again is value weighting. In Fama and French (1998) this happens within countries and across countries, via the MSCI weights. The combined effect again is to give relatively little room to small firms, even less than within the US study: many small firms are classified as locally Big rather than globally Small and then put in the value-weighted global Big portfolio where their impact is minimal. A related consequence of value weighting is that both S and B are now dominated by US firms, making the international sample rather similar to the American one.

Lastly, in Fama & French (1998) there is a requirement that book data be known, and this again eliminates many of the smaller stocks.

We accordingly prefer to construct the size, book-to-market and momentum factor portfolios in one shot, from the global stock list; we use equally weighted portfolio returns for all factors other than the world market; and whenever possible we include also stocks with an unknown book-to-market value.

In the next paragraphs, we investigate whether the alternative composition of the factor portfolios from the US setting (previous section), also applies in an international setting. To do so, we compare the alpha estimates and *t*-statistics of six models for test portfolios, constructed using twelve alternative criteria: (*i*) ten size deciles; (*ii*) ten book-to-market deciles; (*iii*) ten momentum deciles; (*iv*) 25 size-book-to-market portfolios; (*v*) 25 size-momentum portfolios; (*vi*) 25 momentum-book-to-market portfolios; (*vii*) ten ex-US size deciles; (*viii*) ten ex-US book-to-market deciles; (*ix*) ten ex-US momentum deciles; (*x*) 25 ex-US size-book-to-market portfolios; (*xi*) 25 ex-US size-momentum portfolios and (*xii*) 25 ex-US momentum-book-to-market portfolios.

The first model is the basic one-factor CAPM,

$$R_{t+1} - r_{f,t}^{US} = \alpha + \beta\left(R_{m,t+1} - r_{f,t}^{US}\right) + \varepsilon_{t+1}\,. \tag{4}$$

CAPM 2 is Carhart's 4-factor model:

$$R_{t+1} - r_{f,t}^{US} = \alpha + \beta\left(R_{m,t+1} - r_{f,t}^{US}\right) + \gamma SMB_{t+1} + \delta HML_{t+1} + \phi WML_{t+1} + \varepsilon_{t+1}\,. \tag{5}$$

where the factor portfolios are set up following Fama and French (1993, 1995, 1996a, 1996b), Carhart (1997), Jegadeesh and Titman (1993) and Rouwenhorst (1999) except that portfolios are equally weighted and updated monthly.[8]

CAPM 3 is the alternative 6-factor CAPM obtained by adding to (5) the two factors identified in the US tests (the micro-stock and extreme-book-to-market factors, but extracted from the global database) and re-specifying the *SMB*, *HML* and *WML* factors as described in Section 4.2:

$$\begin{aligned}\left[R_{t+1} - r_{f,t}^{US}\right]_i = {}& \alpha_i + \beta_i\left(R_{m,t+1} - r_{f,t}^{US}\right) + \gamma_i mSMB_{t+1} + \delta_i rSMB_{t+1} + \phi_i eHML_{t+1} \\ & + \varphi_i rHML_{t+1} + \theta_i WML_{t+1} + \varepsilon_{t+1}.\end{aligned} \tag{6}$$

The next CAPMs are international ones. The model in Sercu (1980), a static international CAPM that generalizes Solnik (1974), features the world market-portfolio return and the excess returns from investing in each non-USD currency. It has no state variables, so an obvious extension will be to add the standard *SMB*, *HML* and Momentum factors. Including all 39 currencies as factors is not recommendable as the power of the alpha tests will drop dramatically, but apart from this consideration there are no clear guidelines or standard

[8] Thus, *SMB* (small minus big) is the size factor portfolio: a zero-investment portfolio that is long the 50% smallest stocks and short the 50% largest stocks; *HML* (high minus low) is the book-to-market factor portfolio: a zero-investment portfolio that is long the 30% highest B/M stocks and short the 30% lowest B/M stocks; and *WML* (winner minus loser) is the momentum factor portfolio: a zero-investment portfolio that is long the 30% highest past-performers (winners) and short the 30% lowest past-performers (losers).

practices. Jorion (1990) proposes to use a fixed trade-weighted basket of currencies, but this assumes that all stocks have a vector of currency exposures that is proportional to the trade weights, a restriction which Rees and Unni (1999) reject empirically. We adopt a compromise. Specifically, we include in every regression the individual currencies of seven countries (C7), taking at least one currency per continent and looking, per continent, at economic weight and number of stocks in our database. This "C7" list contains the Canadian Dollar, British Pound and Deutsche Mark, Japanese Yen and Korean Won, Australian Dollar and South African Rand. All stocks are allowed to be exposed, without any prior restrictions, to each of these C7 currencies. On top of that, non-C7 stocks are assumed to have a common exposure to their own exchange rate (Adler and Simon, 1986). Since the regressand variables are portfolio returns, this last assumption means that for each such test portfolio a basket of currency deposits is created which gives to each non-C7 currency the same weight as the stocks from that country have in the particular test portfolio. Thus, if a portfolio contains n_1 stocks from non-C7 currency 1 and n_2 stocks from non-C7 currency 2, then the basket consists of fractions $n_1/(n_1+n_2)$ invested in currency 1 and $n_2/(n_1+n_2)$ invested in currency 2. The resulting international version of the 4-factor model reads like

$$\begin{aligned}\left[R_{t+1}-r_{f,t}^{US}\right]_i &= \alpha_i+\beta_i\left(R_{m,t+1}-r_{f,t}^{US}\right)+\gamma_i SMB_{t+1}+\delta_i HML_{t+1}+\phi_i WML_{t+1}\\ &\quad +\sum_{k=1}^{7}\psi_{i,k}XF_{k,t+1}+\zeta_i CXF_{i,t+1}+\varepsilon_{t+1},\end{aligned} \tag{7}$$

$$\text{and } XF_{t+1}=s_{t+1}+r_{f,t}^{*}-r_{f,t}^{US} \text{ with } s_{t+1}=\frac{S_{t+1}-S_t}{S_t}, \tag{8}$$

where the right-side factor portfolios (*SMB*, *HML* and *WML*) are like in Carhart's model. Subscript i stands for the i-th left-side test portfolio, subscript k denotes the k-th exchange factor portfolio and CXF_i refers to the compound non-C7 exchange factor portfolio tailored for the i-th test portfolio. S denotes the going spot exchange rate (USD per foreign currency) and $r^{*}_{f,t}$ the foreign risk-free interest rate. We further refer to this model as the international 4-factor model

The last CAPM candidate is obtained by adding to (7) the two factors identified in the US tests (international versions of the micro-stock and extreme-book-to-market factors) and re-specifying the *SMB*, *HML* and *WML* factors as described in Section 4.2. The resulting international version of the 6-factor model reads like:

$$\left[R_{t+1} - r_{f,t}^{US}\right]_i = \alpha_i + \beta_i\left(R_{m,t+1} - r_{f,t}^{US}\right) + \gamma_i mSMB_{t+1} + \delta_i rSMB_{t+1} + \phi_i eHML_{t+1} + \varphi_i rHML_{t+1} + \theta_i WML_{t+1} + \sum_{k=1}^{7} \psi_{i,k} XF_{k,t+1} + \zeta_i CXF_{i,t+1} + \varepsilon_{t+1}. \tag{9}$$

An equivalent interpretation is that this model adds exchange risk factors to our alternative 6-factor model. We further refer to this model as the international 6-factor model.

5.1 Results for size, book-to-market, and momentum test portfolios

US firms take up 55% of the total sample by numbers, but are relatively underrepresented in the lower size quintile, where they provide only 45% of the observations. Still, when we look at the plot, in Figure 9, of the average monthly dollar return of the ten international size-deciles for the period 1980-1999, it looks a lot like Figure 6 (US market). Thus, also in an international setting there seems to exist a strong small-firm effect for the smallest stocks—unless, of course, the high average return would be explained by beta. In contrast, the inverted small-firm effect in the US market, where the average return of the biggest firms was slightly higher than the average-sized firms, disappears in an international setting: bigger firms earn monotonely less. From Table 7, we see that also in an international setting the specification of the factor portfolios plays an important role in the ability of an asset pricing model to price unmanaged portfolios. Notably, the specification adopted in the US study of the preceding section also produces insignificant alphas for the ten size deciles in an international setting while competing models get rejected.

We now turn to Figure 10 which plots the average monthly dollar return for ten international B/M decile portfolios for the period 1980-1999. Again Figure 10 resembles its US counterpart, Figure 7, which exhibits a gradually rising monthly average return as we move from growth stocks (low B/M value) to book-to-market or value stocks (high B/M value). There is an S-shape, and the highest book-to-market decile sticks out again. From Table 7, we conclude that also in an international setting, the 6-factor models seems to outperform the 4-factor models in pricing unmanaged B/M based test portfolios. Only the third B/M decile portfolio deviates significantly.

Lastly, Figure 11 plots the average monthly dollar return for ten international momentum decile portfolios for the period 1980-1999. Remember that Figure 8 (US market) looked like an S-shaped positive schedule not too far from linearity. Figure 11 resembles a linear rise even more. Apparently, in an international setting, the average return of international momentum portfolios rises at a more constant rate. From Table 7, we see that the Carhart's model does poorly while the 6-factor CAPMs remain superior in pricing those unmanaged international test portfolios.

In the next section we focus on two-dimensionally sorted test portfolios and show that the alternative composition of the risk factors improves the CAPM pricing model considerably.

5.2 Results for two-dimensional test portfolios

In this section we compare the ability of the different models to price two-dimensional size-book-to-market, size-momentum and momentum-book-to-market sorted international portfolios.

In Table 7, the two alternative 6-factor models (one with and one without currency factors) produce no significant alphas against 11, 7 and 5 for the standard models when trying to price 25 size/book-to-market-sorted test portfolios. The hypothesis that all 25 alphas are simultaneously zero cannot be rejected for both alternative 6-factor models at a 5% level of significance. The basic CAPM does worst; it underprices all 5 small stock portfolios and 4 out of 5 book-to-market portfolios. So, the alternative 6-factor models clearly do a better job in pricing 25 unmanaged size-book-to-market sorted portfolios than the standard models.

When pricing test portfolios sorted on size and past performance instead of size and book-to-market, the alternative 6-factor models deliver two significant alphas against 11, 8 and 9 for the standard models. Again the basic CAPM does worst; it overprices 4 out of 5 winner portfolios. So, the alternative risk factor composition significantly improves the ability to price 25 unmanaged size-momentum sorted portfolios.

For the 25 momentum-book-to-market portfolios, the international 6-factor model and the alternative 6-factor CAPM produce, respectively, zero and one significant alpha against 13, 8 and 8 for the standard models. The basic CAPM underprices low-book-to-market losers and overprices high-book-to-market winners. The joint null hypothesis cannot be rejected for the international 6-factor model. Thus, adding exchange risk factors to the alternative 6-factor model does improve the performance, but the effect is small.

5.3 US Impact on the international test

In Section 4.3 we successfully tested the alternative risk factor composition on US data. This success was confirmed on international data in Sections 5.1 and 5.2. In this section we show that the relative large impact of US shares in the international database does not jeopardize the international validation of the alternative risk factor composition. We compute non-US test portfolios by deleting the US stocks from the one- and two-dimensional test portfolios of Section 5.1 and 5.2, preserving the allocation of non-US stocks along the test portfolios. We test the different models on the resulting non-US test portfolios.

From Table 7 we conclude that, even when we exclude the large US effect from the international one-dimensional test portfolios, the basic CAPM keeps on suffering from the well-known anomalies; and the alternative risk factor composition remains superior over the standard composition.

Table 7 shows that in terms of the numbers of significant alphas both 6-factor models and 4-factor CAPMs are equally (un)able to price two-dimensional size-book-to-market sorted non-US portfolios. However the former models produce much smaller individual *t*-statistics and χ^2-statistics. The 4-factor models seem to have difficulties to price small stocks whereas the 6-factor models stumble over small, low-book-to-market stocks stocks. The next row of Table 7 is more clear about the superiority of the 6-factor models over the standard models. The former models have two significant alphas against 11, 8 and 9 for the latter models. Small losers are difficult to price even for the 6-factor models. In the last row of Table 7 the 6-factor models dominate the standard models in terms of significant alphas (zero against 11, 1 and 3), in size of individual *t*-statistics and in χ^2-statistics.

So, even if we exclude the US stocks from the unmanaged size-book-to-market, size-momentum and momentum-book-to-market two-dimensional portfolios, the models with an extra size and book-to-market risk factor do a clearly better pricing job than the standard models.

We conclude that the relative large impact of US stocks on the one- and two-dimensional test portfolios does not invalidate the international evidence of the superiority of the alternative risk factor composition over the standard composition.

6. Conclusion

Prior research appears to have overlooked the impact of smaller stocks by discarding (mostly smaller) stocks for which the database lacks book values or market values. Our database enables several robustness checks. Neither Fama and French's 3-factor nor Carhart's 4-factor model are able to explain the cross-section of stock returns if those smaller stocks get into the picture. We therefore use a new risk factor composition with an extra size and book-to-market factor that models the extreme size and book-to-market risks inherent to those smaller stocks. The resulting alternative 6-factor model is successfully tested against Fama and French's 3-factor and Carhart's 4-factor model with one-dimensional test portfolios sorted on size, book-to-market and momentum, and their two-dimensional combinations even with split samples. To reduce the risk of *ad hoc* modifications, we also test the modification in an international setting of 39 countries, both emerging and developed. We showed that the factors that performed well in the US market also work in an international setting even with US stocks excluded from the test portfolios. Part of the relevance of these results is that they also offer some clues on the causes of the high returns. We identified portfolios that resolve the mispricing via their covariances and narrowed down the list of explanations to factors. This rules out data errors and probably also information asymmetries as explanations of the extra returns.

References

Adler, M., Simon, D., 1986. Exchange risk surprises in international portfolios: some equity markets are super-nominal. Journal of Portfolio Management 12, 44-53.

Amihud, Y., and Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 17, 223-249.

Banz, R. W., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9, 3-18.

Carhart, M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57-82.

Chan, K. C., Chen, Nai-fu, Hsieh, David A., 1985. An exploratory investigation of the firm size effect. Journal of Financial Economics 14(3), 451-471.

Daniel, K., Grinblatt, M, Titman, S., and R. Wermers, 1997. Measuring mutual fund performance with characteristic-based benchmarks, Journal of Finance 52, 1053-1058.

Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. Journal of Financial Economics 7(2), 197-226.

Eun, C., Huang, W., Lai, S., 2003. International diversification with large- and small-cap stocks. Unpublished manuscript.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427-465.

———, 1993. Common factor portfolios in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

———, 1995. Size and book-to-market factors in earnings and returns. Journal of Finance 50, 131-155.

———, 1996a. Multifactor explanations of asset pricing anomalies. Journal of Finance 51, 55-84.

———, 1996b. The CAPM is wanted, dead or alive. Journal of Finance 51, 1947-1958.

———, 1998. Value versus growth: the international evidence. Journal of Finance 53, 1975-1999.

———, 2000. Characteristics, covariances, and average returns: 1929 to 1997. Journal of Finance 55, 389-406.

Jorion, P., 1990. The exchange rate exposure of US multinationals. Journal of Business 63, 331-345.

Jegadeesh, N., Titman, T., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65-91.

Rees, W., Unni, S., 1999. Exchange rate exposure amongst european firms: evidence from France, Germany and the UK. University of Glasgow, AccFin Research Paper 1999/8.

Rosenberg, B., Reid, K., Lanstein, R., 1985. Pervasive evidence of market inefficiency. Journal of Portfolio Management 9, 18-28.

Rouwenhorst, G., 1999. Local return factors and turnover in emerging stock markets. Journal of Finance 54, 1439-1464.

Sercu, P., 1980. A generalization of the international asset pricing model. Revue de l'Association Française de Finance 1(1), 91-135.

Solnik, B., 1974. An equilibrium model of the international capital market. Journal of Economic Theory 8, 500-524.

Stattman, D., 1980. Book values and stock returns. The Chicago MBA: a Journal of Selected Papers 4, 25-45.

Tables

Key to Tables: Boldface (italic) signals significance at a 5% level using SUR standard errors (with White's heteroskedasticity-consistent covariance matrix). "Narrow-based" refers to stocks with both market- and book-value data. "Broad-based" data use all data whenever possible even if book value is missing. "NYSE" or "all" refers to the list—broad or narrow—from which the required deciles are computed

Table 1: Alpha estimates of Fama and French (1996a): original

Size	Book-to-market				
	Low	2	3	4	High
Small	**-0.45**	**-0.16**	-0.05	0.04	0.02
2	-0.07	-0.04	0.09	0.07	0.03
3	-0.08	0.04	0.00	0.06	0.07
4	0.14	**-0.19**	-0.06	0.02	0.06
Big	**0.20**	-0.04	-0.10	-0.08	-0.14

Table 2: Alpha estimates of Fama and French (1996a): replication

Size	Book-to-market				
	Low	2	3	4	High
Small	-0.02	-0.25	0.14	0.18	**0.26**
2	**-0.47**	-0.08	**0.30**	0.14	0.06
3	**-0.33**	0.02	-0.06	0.02	0.02
4	-0.09	-0.01	-0.10	-0.07	0.00
Big	**0.19**	-0.01	0.03	-0.07	**-0.39**

Table 3: Broad-based factor portfolios and breakpoints (all); broad-based test portfolios

Size	Book-to-market				
	Low	2	3	4	High
Small	**0.40**	**0.42**	***0.38***	-0.03	***0.62***
2	*-0.47*	**-0.40**	***-0.55***	***-0.62***	***-0.48***
3	-0.32	-0.15	***-0.51***	***-0.37***	-0.15
4	0.20	-0.17	-0.32	-0.23	-0.27
Big	*0.30*	***-0.25***	***-0.41***	**-0.47**	0.03

Table 4: Summary of Test Design Variations

	Test Design Variations	**# Significant $\hat{\alpha}$'s**	$\mathrm{E}(\hat{\alpha})$	$\mathrm{E}(\lvert\hat{\alpha}\rvert)$
1a	Starting point (Table 2): Fama and French (1996a) replication	6	-0.02	0.13
1b	Time period's alternative: 1976-1993 *to* 1994-1999	7	-0.17	0.31
1c	Time period's alternative: 1976-1993 *to* 1980-1999	6	-0.07	0.17
2	R_m & R_f: CRSP *to* Datastream & T-bill *to* discount rate	5	-0.09	0.18
3a	Update frequency *test* portfolios: yearly *to* monthly	13	-0.02	0.30
3b	Update frequency *factor* portfolios: yearly *to* monthly	13	-0.19	0.23
3c	Update frequency *test- & factor* portfolios: yearly *to* monthly	14	-0.19	0.27
4a	Weighting scheme *test* portfolios: value *to* equal	14	-0.17	0.28
4b	Weighting scheme *factor* portfolios: value *to* equal	15	-0.31	0.35
4c	Weighting scheme *test- & factor* portfolios: value *to* equal	16	-0.29	0.34
5	Broad-based *factor* portfolios	9	-0.10	0.24
6	Broad-based *breakpoints (NYSE)*	7	-0.05	0.25
7	Broad-based *test* portfolios	14	-0.23	0.29
8	Broad-based *breakpoints (all)*	13	-0.15	0.34
9	Ending point (Table 3): OLS' *to* White's errors	11	-0.15	0.34

Key to Table 4:
- Row 1a, 1b and 1c only differ regarding the years of data, with Row 1a showing the pre-1994 alphas (the overlap with Fama and French, 1996a), Row 1b displaying the post-1993 results, and Row 1c the number of significant alphas for the full sample.
- From Row 1c to Row 2 the risk-free rate becomes the US discount rate instead of the US T-bill rate,[12] and the market return is Datastream's US market return, not the value-weighted return on all stocks in the size-book-to-market portfolios plus the negative-book-value equities as in Fama and French (1993, 1996a).
- In the data underlying Row 3c, the compositions of both the test and the factor portfolios are updated every month instead of yearly. We introduce monthly updating in turn on the test and factor sides. Thus, the changes between Row 3a and Row 3b are not cumulative.
- We introduce equal weights in turn on the test and factor sides. Thus, the changes between Row 4a and Row 4b are not cumulative.
- In Rows 5 to 9, the factor portfolios are built from all stocks, not just those with both market- and book-value data. Specifically, the size factor is now computed as the difference between the equally weighted average return for all stocks above the median versus the average for all below-median stocks, whether they provide book-value information or not; similarly, the book-to-market factor is the difference of the equally-weighted returns on portfolios containing the firms that rank below the 30th or above the 70th percentile re B/M. In this new test, the coverage for book-to-market is the same as before, since market values are almost never missing. The most powerful results (in the sense of providing the highest number of rejections) were obtained as follows. We keep the earlier 25 pure-intersection portfolios as the starting basis of the new test

[12] For reasons of availability and international comparability our risk-free rates are from the IMF's *International Financial Statistics*. For the US T-bill, this source provides only a monthly average of the 3-month rate. The IMF US discount rate, in contrast, is an end-of-period rate and a similar number is available for most countries. Eligible depository institutions pay this rate when borrowing short-term from a Federal Reserve Bank.

portfolios. The additional stocks, those with just size information, are sorted into the five size buckets, and from there are transferred to one of the 25 old intersection portfolios, taking care to stay within the same size bracket but randomizing across B/M category. This procedure shrinks the dispersion across book-to-market classes, but everything else being the same, also reduces the noise in the portfolio returns.[13] Thus, whether on balance power improves or not is an empirical matter.

- In row 8 and 9, we calculate the breakpoint values on the quintile values from the entire dataset, not just the NYSE ones.[14] Again, the size coverage of the portfolios widens because the number of assets per size or B/M group is now equal across groups rather than very much bunched together at the small-cap or high-growth end. The effect of the new way of defining the buckets becomes stronger over time, this time: in 1980 the non-NYSE list in Datastream represents just 13% of the total, but that percentage rises to over 70% in 1999. Non-NYSE firms in Datastream had a mean market value of less than one-fourth of the typical Big-Board listee in 1980, and about 45% in 1999. Thus, as expected, computing the quintiles from the all-stock list brings about drastically lower quintile values for especially the first quintiles.

Table 5: Number of significant alphas and Wald's *p*-value: one-dimensional test portfolios

test portfolios sorted on (10)	**Carhart's 4-factor model**	**Carhart: SMB & HML: updated yearly**	**Carhart: R_m: leads & lags**	**Carhart: SMB & HML: Fama & French**	**Alternative 6-factor model**
Size	1 (0.00)	1 (0.00)	2 (0.00)	2 (0.00)	0 (0.88)
B/M	3 (0.00)	3 (0.00)	3 (0.00)	5 (0.00)	0 (0.30)
Momentum	2 (0.00)	0 (0.00)	2 (0.00)	6 (0.00)	0 (0.37)

Key to Tables 5, 6 and 7: The construction of the factor and test portfolios always proceeds as follows. Decile breakpoints are set using all stocks, including Amex and NASDAQ firms. Stocks with an unknown market value or book-to-market value are also used to set breakpoints and to calculate the factor and test portfolios (i.e. broad-based factor portfolios and breakpoints (all); broad-based test portfolios), and all portfolios are equally weighted and updated monthly. All regression test *t*-statistics are computed using a SUR specification that accounts for intertemporal heteroskedasticity within each series beside, of course, cross-equation heteroskedasticity and correlation. SMB (small minus big) is the size factor portfolio, *viz*, a zero-investment portfolio that is long the 50% smallest stocks and short the 50% largest stocks. HML (high minus low) is the book-to-market factor portfolio, a zero-investment portfolio that is long the 30% highest B/M stocks and short the 30% lowest B/M stocks. WML follows the Rouwenhorst (1999) version of Jegadeesh and Titman (1993). Stocks are ranked on the basis of the return realized in the months t-7 to t-2. (Month t-1 is omitted to eliminate the common bid-ask-bounce effect that would otherwise have affected both the past performance and the subsequent return.) All available data are used, whether book value is available or not. The momentum factor is the equally-weighted return, for month *t*, on the 30% best winners minus the 30% worst losers. Contrary to the size- and book-to-market-portfolios, momentum-portfolios use a holding period of not one month but six, as in Rouwenhorst (1999) or Jegadeesh and Titman (1993). Again following these authors, we compute the monthly average return across the six ongoing momentum

[13] As an alternative, we tried working with unions rather than intersections. Under that procedure we allocate every stock with a known market value into one of five size groups and all stocks with a known book-to-market into one of five book-to-market groups. This gives us ten basic test portfolios. We then form a portfolio for size-book-to-market combination (i, j) as the average of size portfolio i and book-to-market portfolio j, weighted by the number of stocks in i and j, respectively, thus computing 25 different combinations of the ten basic alphas, similar to the 25-portfolio tests used thus far. The outcome was a somewhat larger number of rejections (nine, up from seven) despite generally lower alphas—a signal of higher power relative to the original Fama and French (1993) design but nowhere as strong as the alternative procedure outlined in the main text. In addition, this induces strong dependencies across test portfolios.

[14] Besides allowing more attention to small stocks, another consideration for basing the breakpoints on the entire sample rather than the list of the leading exchange is that the latter procedure cannot be applied consistently across countries.

strategies, each started one month apart, to handle the issue of overlapping observations. Like the size and book-to-market portfolios, the momentum portfolios are updated monthly and are equally weighted.

Table 6: Number of significant alphas and Wald's χ^2-value: two-dimensional test portfolios

Test portfolios sorted on (25)	**Carhart's 4-factor model**	**Alternative 6-factor model**
Size & B/M	7 (173.12)	1 (53.70)
Size & Momentum	4 (160.68)	1 (42.75)
Momentum & B/M	4 (106.26)	0 (54.53)
Size & B/M: separate data	5 (130.76)	1 (57.02)

Key to Table 6: Fama and French (1995) point out that spurious common variation might be induced when the regressor portfolios SMB and HML are constructed from the same stocks as the regressand test portfolios. To avoid this, stocks are ranked alphabetically; the odd-numbered are used to calculate the left-side test portfolios, from the even-numbered stocks we calculate the right-side risk factors (i.e. separate data).

Table 7: Number of significant alphas and Wald's χ^2-value: international portfolios

Test portfolios sorted on	**Basic 1-factor**	**Carhart's 4-factor**	**Alternative 6-factor**	**International 4-factor**	**International 6-factor**
Size (10)	1 (423)	3 (324)	0 (16.23)	4 (311)	0 (13.81)
B/M (10)	4 (61.60)	3 (37.33)	1 (18.51)	3 (52.91)	1 (19.31)
Momentum (10)	4 (29.10)	2 (32.30)	0 (39.36)	2 (25.22)	0 (29.23)
Size & B/M (25)	11 (228)	7 (177)	0 (35.99)	5 (166)	0 (35.62)
Size & Momentum (25)	11 (142)	8 (224)	2 (112)	9 (208)	2 (94.49)
Momentum & B/M	13 (145)	8 (83.72)	1 (48.30)	8 (76.26)	0 (29.92)
Non-US Size (10)	4 (201)	1 (137)	0 (52.53)	1 (142)	2 (54.80)
Non-US B/M (10)	2 (43.01)	2 (47.89)	0 (18.67)	3 (52.39)	0 (18.54)
Non-US Momentum (10)	2 (21.24)	0 (18.31)	0 (9.51)	0 (16.39)	0 (7.60)
Non-US Size & B/M (25)	11 (145)	5 (138)	4 (66.65)	4 (111)	5 (58.44)
Non-US Size & Momentum (25)	11 (142)	8 (224)	2 (112)	9 (208)	2 (94.49)
Non-US Momentum & B/M	11(95.41)	1 (69.29)	0 (39.23)	3 (67.92)	0 (35.79)

Key to Table 7: Non-US test portfolios are composed by deleting the US stocks from the original one- and two-dimensional test portfolios preserving the allocation of non-US stocks along the test portfolios

Figures

Figure 1: Geographical distribution of the stock list (in %)

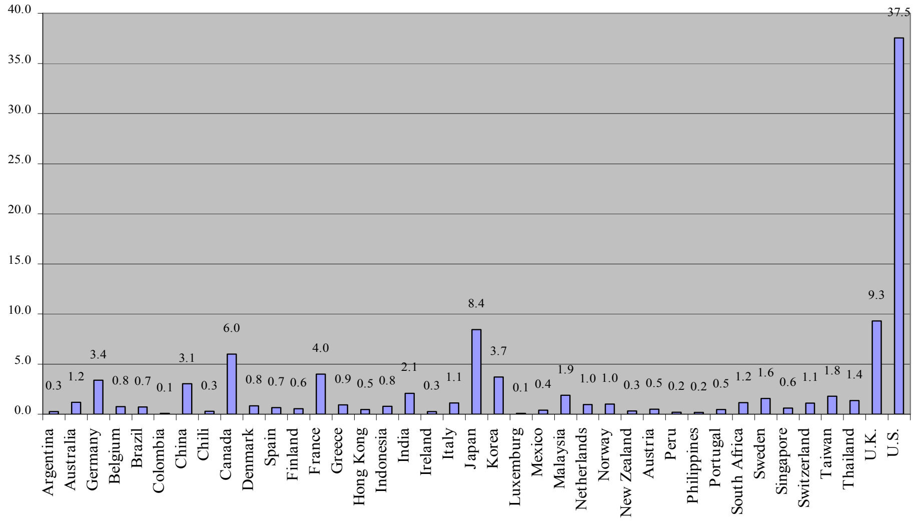

Figure 2: Equally weighted country indices: mean monthly return

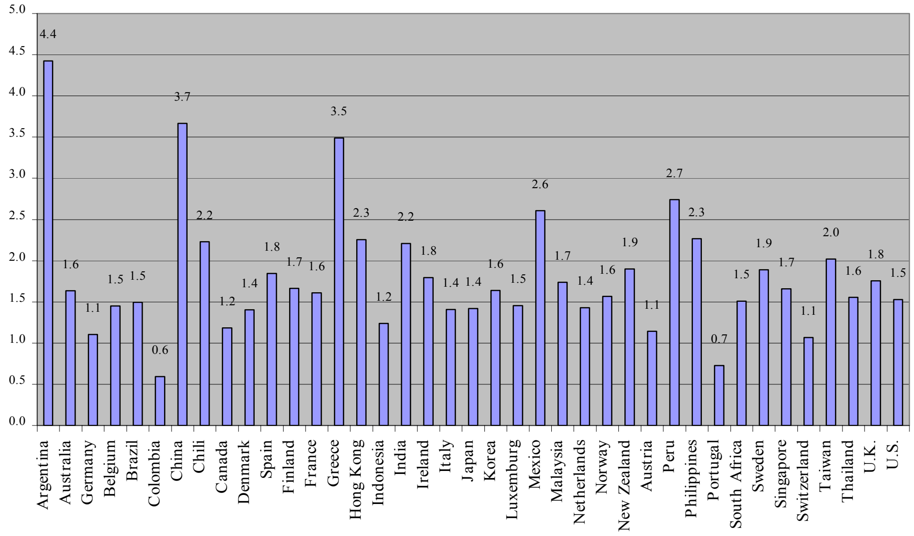

Figure 3: Equally weighted country indices: standard deviation of monthly return

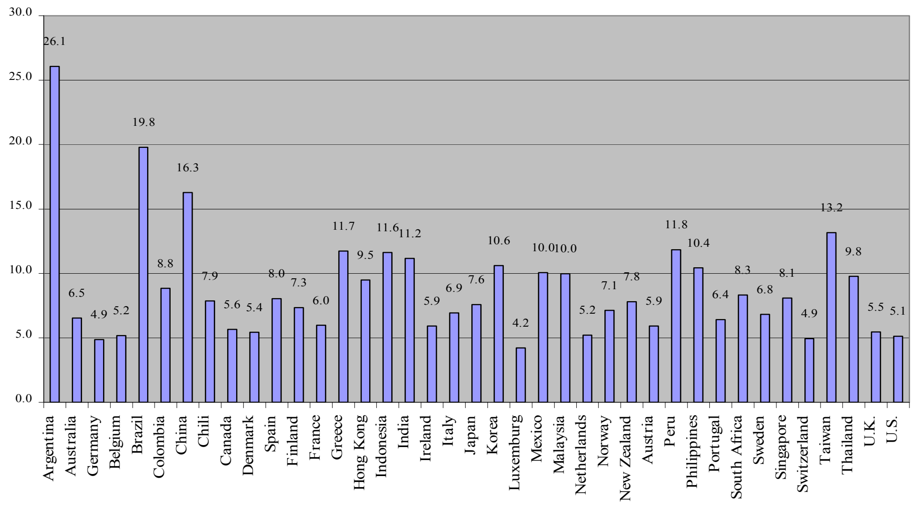

Figure 4: Equally weighted country indices: skewness of monthly return

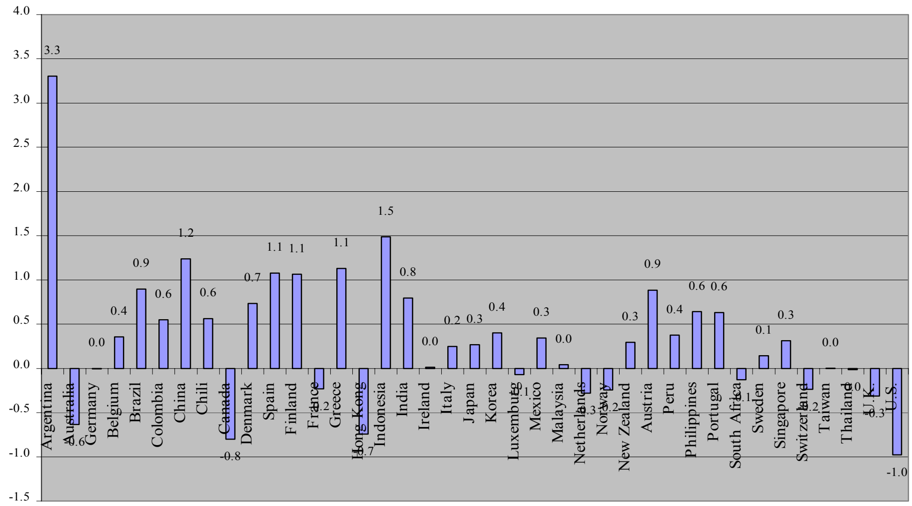

Figure 5: Equally weighted country indices: excess kurtosis of monthly return

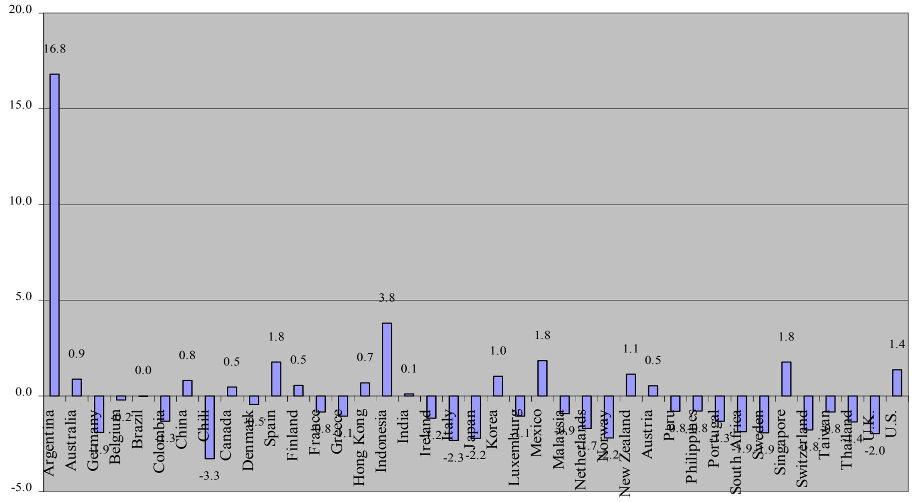

Figure 6: $\overline{R}$ for 10 size deciles: US data

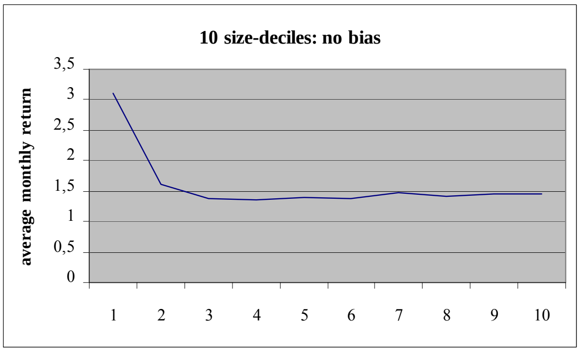

Figure 7: $\overline{R}$ for 10 B/M deciles: US data

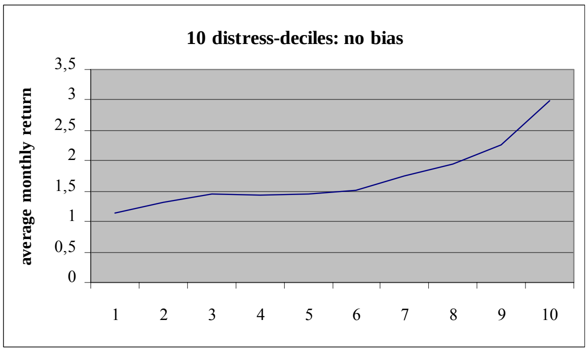

Figure 8: $\overline{R}$ for 10 MoM deciles: US data

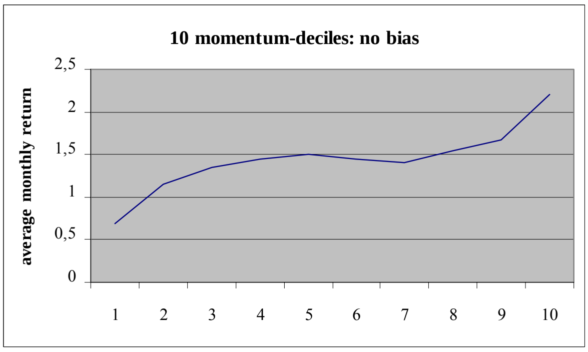

Figure 9: $\overline{R}$ for 10 size deciles: Int. data

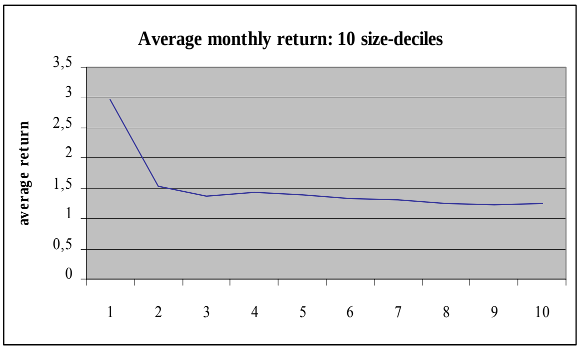

Figure 10: $\overline{R}$ for 10 B/M deciles: Int. data

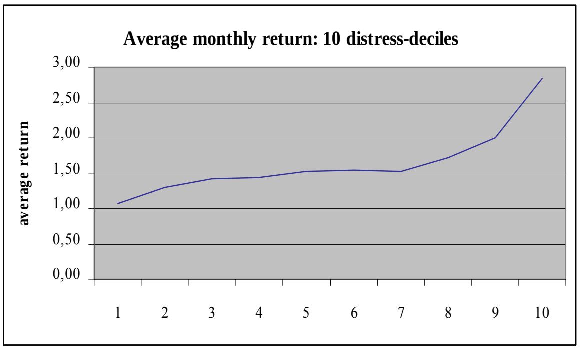

Figure 11: $\overline{R}$ for 10 MoM deciles: Int. data